Investigating yellow dung fly body size evolution in the field: Response to climate change?

Blanckenhorn, Wolf U

Abstract: Uncovering genetic responses to selection in wild populations typically requires tracking individuals over generations and use of animal models. Our group monitored the body size of one Swiss Yellow Dung Fly (Scathophaga stercoraria; Diptera: Scathophagidae) field population over 15 years, including intermittent common-garden rearing in the laboratory to assess body size with minimized environmental and maximized genetic variation. Contrary to expectations based on repeated heritability and phenotypic selection assessments over the years (reported elsewhere), field body sizes declined by >10% and common-garden laboratory sizes by >5% from 1993 to 2009. Our results confirm the temperature-size rule (smaller when warmer) and, albeit entirely correlational, could be mediated by climate change, as over this period mean temperature at the site increased by 0.5°C, although alternative systematic environmental changes cannot be entirely excluded. Monitoring genetic responses to selection in wild invertebrate populations is thus possible, though indirect, and wild populations may evolve in directions not consistent with strongly positive directional selection favoring large body size.

DOI: https://doi.org/10.1111/evo.12726

INVESTIGATING YELLOW DUNG FLY BODY SIZE EVOLUTION IN THE FIELD: RESPONSE TO CLIMATE CHANGE?

Wolf U. Blanckenhorn

Institute of Evolutionary Biology and Environmental Studies, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland; Fax: +41 44 635.4780; E-mail: wolf.blanckenhorn@ieu.uzh.ch

For *Evolution* as a *Short Communication*

Word count Abstract 163; Text 4310; 47 References 1179; 1 Figure; 2 supplementary Tables

Abstract Uncovering genetic responses to selection in wild populations typically requires tracking individuals over generations and use of animal models. Our group monitored the body size of one Swiss Yellow Dung Fly (*Scathophaga stercoraria*; Diptera: Scathophagidae) field population over 15 years, including intermittent common garden rearing in the laboratory to assess body size with minimized environmental and maximized genetic variation. Contrary to expectations based on repeated heritability and phenotypic selection assessments over the years (reported elsewhere), field body sizes declined by >10% and common-garden laboratory sizes by >5% from 1993 – 2009. Our results confirm the temperature-size rule (smaller when warmer) and, albeit entirely correlational, could be mediated by climate change, as over this period mean temperature at the site increased by 0.5 °C (ca. 5%), although alternative systematic environmental changes cannot be entirely excluded. Monitoring genetic responses to selection in wild invertebrate populations is thus possible, though indirect, and wild populations may evolve in directions not consistent with strongly positive directional selection favoring large body size.

KEY WORDS: Body size evolution, climate change, common garden experiment, quantitative genetics, *Scathophaga stercoraria*, selection, temperature

INTRODUCTION

The evolution of body size continues to captivate biologists, as this trait affects most aspects of the physiology, ecology, life history, and fitness of organisms. Fecundity selection favors large females and sexual selection large males in many species (Roff 1992; Andersson 1994; Kingsolver and Pfennig 2004). In contrast, the counterbalancing forces favoring small body size, presumably primarily caused by juvenile viability costs of becoming large and/or costs of maintaining a large adult size (Schluter et al. 1991), are poorly documented, though there also are a number viability benefits of being large (Blanckenhorn 2000). Reasons why the latter remain enigmatic include lack of research and technical difficulties with detecting size-dependent mortality in the wild (Blanckenhorn 2000). Given generally substantial trait heritability h^2 (Mousseau and Roff 1987), predominant fitness advantages of large body size in combination with few disadvantages should lead to increasing body size in wild populations over evolutionary time in response (R) to such net positive directional selection S: a within-species, micro-evolutionary version of Cope's rule (Kingsolver and Pfennig 2004). Yet, it is generally difficult to detect such micro-evolutionary size increases predicted from the breeder's equation ($R = h^2 * S$) when studying wild populations (Merilä et al. 2001), rendering evolutionary interpretations based on selection estimates and those based on population responses discordant (Gotanda et al. 2015).

Identification of evolutionary responses in nature first and foremost requires time-consuming but very valuable long-term studies of wild populations spanning multiple generations. Such studies of free-ranging animals in their natural environment exist particularly for larger vertebrates, which can be individually identified, marked and followed throughout their lives as long as they remain in a defined area. Examples include mammals like red deer (Nussey et al. 2008), wild sheep (Réale et al. 1999; Coltman et al. 2001, Wilson et al. 2007) or squirrels (Réale et al. 2003), birds like great or blue tits (Charmantier et al. 2004; Husby et al.

2011), mute swans (Charmantier et al. 2006), or house sparrows (Jensen et al. 2003), and even sharks (DiBattista et al. 2007). Repeated observation of identifiable individuals further permits tracking of genetic pedigrees over many generations in the wild, and thus assessment of genetic parameters such as heritabilities of and genetic correlations among various life history traits by employing so-called statistical 'animal models' (Kruuk 2004). Thus not only phenotypic but also genetic responses can be tracked (Merilä et al. 2001; Kruuk 2004; Merilä and Hendry 2014), which brings us the crucial step closer to investigating evolutionary responses to selection in nature. Long-term, cross-generational field studies are essentially impossible to perform with small invertebrates, which cannot be easily followed in nature. Repeated common garden rearing in the laboratory via more traditional parent-offspring or sib analyses (e.g. Blanckenhorn 2002) is an alternative indirect method permitting inferences about genetic changes over time (Stoks et al. 2014). Here I use this approach to study body size evolution of a wild yellow dung fly (*Scathophaga stercoraria*; Diptera: Scathophagidae) population over a 15-year period spanning roughly 60 (overlapping) generations.

Our group worked with one field population of yellow dung flies in Fehraltorf (N47˚23', E8˚44') near Zürich, Switzerland, from 1993 to 2009; thereafter the farmer altered his livestock management and no longer kept the cows outside. During our studies, which are summarized in Blanckenhorn (2007, 2009) and Blanckenhorn et al. (2010), we multiply documented strong and consistent sexual and fecundity selection favoring large male and female body size in this species (summarized in table 10.1 in Blanckenhorn 2007). Directional sexual selection coefficient averages favoring large male size estimated in the field over several years ranged from very high S_{Sex} = +0.275 to +0.505 (cf. Hoekstra et al. 2001; Kingsolver et al. 2001), and average fecundity selection differentials favoring large female size from S_{Fec} = +0.187 to +0.223 (Jann et al. 2000; Kraushaar and Blanckenhorn 2002; Blanckenhorn et al. 2003). Although Teuschl et al. (2007) also demonstrated some counter-selection in terms of seasonal and winter

juvenile viability favoring short development, and hence smaller body size, on the order of S_{Via} = -0.014 to -0.311, this counter-selection does not match positive selection on male or female body size in yellow dung flies (Blanckenhorn 2007).

As is typical, adult body size in this species is a highly plastic trait mediated by variation in juvenile growth rate and development time, and all these traits strongly depend on environmental conditions such as food (here dung) availability and quality, temperature, photoperiod, or predators encountered during development. For example, Amano (1983) reared yellow dung flies at ever increasing dung availabilities, showing that they reached an asymptotic maximal body size beyond 2 g of fresh dung per individual. Selection acts on the phenotype (and thus only indirectly on the genotype), so an overwhelming environmental component of body size could in principle explain any lack of a genetic evolutionary response (Merilä et al. 2001). Over the years we repeatedly sampled and measured field-caught male and female individuals from our population, often multiple times over the season. At the same time, we regularly reared grand-offspring of field-caught flies in the laboratory at various temperatures and food conditions. Such common garden rearing at (over)abundant, "ideal" conditions at least standardizes if not eliminates, but certainly minimizes most extraneous and unexplained environmental body size variation present in the field, thus exposing and maximizing the heritable component (Weigensberg and Roff 1996). We have also multiply estimated the broad- and narrow-sense heritability h^2 (the latter based on additive genetic variance) of body size in our yellow dung fly population to be at least 35%, using various methods (parent-offspring or sib analyses; realized heritability after artificial selection by Teuschl et al. 2007; data and methods summarized in Blanckenhorn 2002). According to the breeder's equation, $R = h^2 * S$, I therefore expected the heritable body size of yellow dung flies to respond to net positive directional selection by increasing over time. Here I analyze body size evolution in our yellow dung fly population over 15 years based on field as well as laboratory-reared flies.

METHODS

Study species

Yellow dung flies abound throughout the northern hemisphere and are particularly common around cow pastures in north-central Europe. They prefer cooler climates and occur at high latitude sites such as Spitzbergen, Greenland or Iceland. In the warmer, southern areas of its distribution, the species inhabits higher altitudes like the Pyrenees or the Sierra Nevada in Spain. In lowland Switzerland and other European countries, each year has a spring (March – June) and an autumn season (September – November), while during the hot midsummer (July and August) the flies largely disappear from the pastures due to their heat sensitivity (Blanckenhorn 2009; Blanckenhorn et al. 2010).

Yellow dung fly adults are sit-and-wait predators of smaller flying insects, requiring an external protein source to develop sperm and eggs (i.e. they are anautogenous income breeders). Females spend most of their time foraging for nectar (for energy) and prey (for reproduction) in the vegetation surrounding the pastures. About once a week they lay a clutch of eggs into fresh dung of large mammals, mainly cattle, on which the larvae feed and wherein they develop. Males wait at the dung pat for incoming females to mate with, so there flies can be caught in high numbers using nets or small vials, which we have done repeatedly over the years (summarized in Table S1). Larvae face unpredictable spatio-temporal variation in dung (i.e. food) quality and quantity, intra- and inter-specific competition, and dung drying, all potentially reducing their phenotypic adult body size. Towards the end of the season the flies have to reach the overwintering pupal stage before the first winter frost (Blanckenhorn 2009; Blanckenhorn et al. 2010).

General laboratory rearing procedures

In the laboratory, field-caught or laboratory-reared flies were allowed to copulate in small bottles equipped with a standardized portion (ca. 5 g) of defrosted cow dung with a genetically unrelated male. Depending on their body size and nutrition, females lay clutches of 30 to 90 eggs partially submerged into the dung. For propagation of the next generation, as many eggs as desired are then transferred with curved forceps into disposable plastic containers with excess fresh cow dung (> 2 g per larva: Amano 1983). Excess dung ensures little (i.e. no) intra-specific competition at overabundant food, and in combination with typically constant laboratory rearing conditions (temperature, humidity, etc.) minimizes the environmental and thus maximizes the genetic component of the phenotypic trait variance (here body size: Weigensberg and Roff 1996; Roff 1997). Depending on temperature (see Table 1 in Blanckenhorn et al. 2010), adult flies emerge after a minimum of 17 days pre-adult development time.

Our common garden laboratory studies conducted over the years in about half of the cases involved quantitative genetic trait assessments in common environments (Table S2) using flies derived from large and outbred parental base populations of ≥ 30 females randomly collected at Fehraltorf. Parent-offspring and/or sib analyses were generally conducted with F2 (rarely F1, and sometimes F3; see Table S2) laboratory generation grand-offspring to eliminate any carry-over maternal effects (methods detailed in Supplement 1). However, this is of minor relevance here, as I focused on common garden studies to estimate phenotypic body size means in the laboratory that were conducted at various constant rearing temperatures ranging from 12 °C to 26 °C, most commonly 18 °C and 24 °C (Table S2). I avoided (i.e. excluded) food and any other manipulations that limit body size (cf. Amano 1983). That is, per capita food = dung supply and larval competition were always controlled or standardized experimentally, whereas rearing temperature was here controlled statistically by including it in the model, as described below. In general, we split maternal (full-sib, sometimes half-sib) broods into various temperature environments. Humidity was always set at 60%. Dung quality, another important

but elusive factor known to influence the phenotype, was standardized within studies by dung mixing and subsequent freezing, but not across studies. While this factor certainly introduces unexplained variation (Blanckenhorn et al. 2010), there is little reason to suspect systematic changes of cow dung over the years, which however cannot be excluded with certainty.

We always measured hind tibia length as an index of size, which is highly repeatable within samples when measured by one person ($R \approx 0.98$), to the extent that this allows assessment of fluctuating asymmetry (e.g. Blanckenhorn et al. 2003). Repeatability across samples could not be verified because samples were measured in the year of the study by different people with different methods, with or without various software, and were not kept until the end to be re-measured, which would be prohibitive given a total sample of >20000 flies. Again, however, there is no reason to suspect systematic biases over the years, though this admittedly cannot be excluded with certainty. Consult Supplement 1, Blanckenhorn (2002, 2009) and Blanckenhorn et al. (2010) for further methodological details.

Statistical analyses

I analyzed field and laboratory common garden body sizes similarly using ANCOVA in SPSS V.21. Temporal sample means weighted by sample size as given in Tables S1 and S2 were used as independent data points, as there are ca. four overlapping generations per year and individuals can only be caught and measured once. Sampling time (year.month, in decimals, for field sizes, and year for laboratory sizes) and temperature (mean monthly averages from the nearest weather station (Kloten; ca. 15 km away) measured in a given year for the field, and constant rearing temperature for the laboratory) were included as continuous covariates (as given in Tables S1 and S2), together with sex as a *repeated* grouping factor: Y (body size) = μ + sex + (rearing or measured) temperature + time (i.e. year(.month)) + sex*time + sex*temperature + ε. This analysis acknowledges that male and female body sizes of a given sample are not completely

independent, which is obvious for the common garden laboratory sizes because males and females were siblings in all cases, but is also the case to a lesser extent for the field samples, if only because males and females grew up under similar environmental conditions. I initially also included quadratic covariate terms, which however were removed because not significant. Sex by covariate interactions were left in the model even if non-significant. To verify the sex-specific regression slope estimates, I additionally performed separate analyses for males and females. My analysis effectively bases the significance tests and estimates on the (weighted) mean body sizes of both sexes, overall sample size thus being equal to the number of temporal samples.

RESULTS

Figure 1a shows the (raw, i.e. uncorrected) field and Figure 1b the common garden laboratory body size means for male and female yellow dung flies from Fehraltorf from 1993 to 2009. The field sample is based on a total of ca. 4000 field-caught and measured females and 11500 males, and the laboratory sample on ca. 3500 laboratory-reared females and 4000 males. Sample sizes of the individual point sample means plotted in Fig. 1 range from 15 to 351 (102 on average) for females and 28 to 1101 (274 on average) for males in the field, and 22 to 380 (121 on average) for females and 22 to 736 (116 on average) for males in the laboratory (Tables S1 and S2). Temporal body size variation of the flies at Fehraltorf should thereby have been estimated with good precision.

AIC analysis indicated that the quadratic fit is no better than the linear fit, so I only report the latter. Field males were overall larger than females (3.20 mm ± 0.03 vs. 2.42 ± 0.03), as is the norm in this species (Blanckenhorn et al. 2010; sex effect: $F_{1,34} = 11.09$, $P = 0.002$). Crucially, field mean body sizes of both sexes declined over the 15 years by more than 10% (overall time effect: $F_{1,34} = 29.28$, $P < 0.001$; Fig. 1a), more steeply so in males than females (sex by time interaction: $F_{1,34} = 10.61$, $P = 0.003$). Male hind tibia length decreased by b = -

0.034 ± 0.006 (SE) mm/year (partial r = -0.70) from 3.47 mm to 2.96 mm (total of 0.51 mm, or 14.8%), while female hind tibia length decreased by b = -0.017 ± 0.005 mm/year (partial r = -0.52) from 2.56 mm to 2.30 mm (total of 0.255 mm, or 10.0%; Fig. 1a). Taking these values as the response R to selection on body size and conservatively assuming a body size heritability of h^2 = 0.35 (Blanckenhorn 2002; Teuschl et al. 2007), the strength of selection of unknown nature can be estimated from the breeder's equation as $S = R / h^2$, yielding S_{male} = -0.098 mm/year for males and S_{female} = -0.049 mm/year for females. Corresponding mean monthly field temperatures from the nearest weather station also correlated negatively with mean body size in this multiple regression, thus apparently additionally explaining some of the residual variation in body size within and among years in agreement with the temperature-size rule (see below), although inclusion of this variable overall did not significantly improve the model beyond the overall effects on hind tibia length already encompassed by sampling time (males: partial r = -0.21, b = -0.008 ± 0.006 mm/°C, P = 0.188; females: partial r = -0.27, b = -0.009 ± 0.005 mm/°C, P = 0.077; sex by temperature interaction: P = 0.873). Analogous sex-specific analyses, including those few samples with missing data for the other sex (Table S1) that were excluded in the repeated-measures analyses, yielded virtually identical regression estimates.

Analogous multiple linear regression (again no better quadratic fit, hence dropped) of mean laboratory body sizes under common garden conditions as a function of time (= year in this case) and constant rearing temperature (see Table S2) showed that laboratory flies also became smaller by ca. 6% over the same 15-year period (time = year effect: $F_{1,25}$ = 19.53, P < 0.001, when only considering the subsample with both male and female data in Table S2; Fig. 1b). Males were again larger than females (3.64 mm ± 0.02 vs. 2.90 ± 0.02; sex effect: $F_{1,25}$ = 45.92, P < 0.001), though here the sex by time interaction was not significant ($F_{1,25}$ = 0.10, P = 0.761). (Laboratory flies were also larger than field flies, but this was not specifically tested.) Male laboratory hind tibia length decreased by b = -0.015 ± 0.004 mm/year (partial r = -0.39) from

3.73 mm to 3.55 mm (total of 0.23 mm, or 6.2%), and female hind tibia length decreased by b = -0.014 ± 0.003 mm/year (partial r = -0.55) from 2.97 mm to 2.80 mm (total of 0.21 mm, or 7.1%). At the same time, and as expected and previously known (Blanckenhorn et al. 2010), flies also grew smaller at warmer rearing temperatures ($F_{1,25}$ = 42.33, P < 0.001; males: partial r = -0.65, b = -0.032 ± 0.006 mm/°C; females: partial r = -0.53, b = -0.015 ± 0.005 mm/°C; sex by temperature interaction also significant: $F_{1,25}$ = 8.92, P = 0.006) in accordance with the temperature-size rule (Atkinson and Sibly 1997). Again, the sex-specific analyses including the few samples with missing data for the other sex (Table S2) yielded virtually identical regression estimates.

Based on temperature records of the Swiss Meteorological Association (SMA) for the nearest site, mean *annual* temperature (also given in Table S1) on average increased by ca. 0.5 °C (~ 5%) over the sampled 15 years (1994 – 2009; b = 0.032 ± 0.028). Again, the quadratic fit did not significantly improve the model.

DISCUSSION

Even though it is practically impossible to conduct multi-generational field studies with small invertebrates that cannot be tracked over their lifetime in similar ways as larger vertebrates (Réale et al. 1999, 2003; Coltman et al. 2001; Jensen et al. 2003; Charmantier et al. 2004, 2006; DiBattista et al. 2007; Nussey et al. 2008; Husby et al. 2011), I here implemented a crude approach to the problem in form of a miniature meta-analysis of our own accumulated data. I investigated evolutionary responses over time in the wild by repeatedly assessing the body size of grand-offspring of field-collected flies reared under non-limiting common garden laboratory conditions, thus minimizing unexplained environmental size variation and exposing the maximal heritable component (Weigensberg and Roff 1996). The main objective of this study was proof of such principle, to help offset especially recent criticism that studies of phenotypic selection do

not directly assess evolutionary responses and can therefore be misleading (Kingsolver and Pfennig 2004; Merilä and Hendry 2014; Gotanda et al. 2015). My approach addresses and likely deflects at least two potential reasons for lacking responses to selection in the wild: that the genetic response is present but masked by a changing environment, and/or that selection acts primarily on the environmental rather than the genetic component (Merilä et al. 2001). I found that the common garden (laboratory) body size change over 15 years in yellow dung flies (Fig. 1b) was roughly 50% of the phenotypic size change estimated in the wild (Fig. 1a). As the heritability of body size estimated for this species is at least 35% (Blanckenhorn 2002; Teuschl et al. 2007), this suggests that selection, of whatever and unknown nature, in the wild has affected environmental and genetic components roughly equally, although this of course cannot be inferred with certainty from the present phenotypic data. Regardless, I encourage researchers to apply this approach to other invertebrates (cf. Schilthuizen and Kellermann 2014; Stoks et al. 2014).

The second objective of my study was to explain the results obtained. Wild yellow dung flies of our study population in Fehraltorf near Zurich have decreased in body size (measured as hind tibia length, a standard index for size in this and other insect species: Blanckenhorn 2009) by more than 10% from 1994 to 2009 (Fig. 1a). This size change was not merely phenotypic, but was matched by their common garden size estimated throughout the same time period, which also decreased by about 6% (Fig. 1b). Weighting by sample size slightly affects the estimates but not the overall result. As my study is phenomenological and correlational, I can only speculate about the causes, but the average temperature at our site increased by ca. 0.5 °C over the studied time period. Although adding mean monthly field temperatures for the respective years to the regression of field body size on sampling time (Fig. 1a) did not significantly improve the model, body size did also decrease slightly and non-significantly with temperature independent of time, thus additionally explaining a bit of the residual body size deviations in the expected direction in

accordance with the temperature-size rule (smaller when warmer: Atkinson and Sibly 1997). It should be obvious that the sampling time (i.e. month and year) effect already conceptually incorporates a great portion of any variation caused by medium-term climatic changes, such that any collinear effects of time and temperature ultimately cannot be statistically separated in principle.

The temperature-size rule is a common pattern in ectothermic animals related but not identical to Bergmann's rule (smaller at lower latitudes: Blanckenhorn and Demont 2004; Chown and Gaston 2010) that describes or predicts a body size decrease with temperature. My crude analysis here based on laboratory common garden rearing at various temperatures recovered this temperature-size effect that has been documented multiply before in this and other species (Blanckenhorn 1997, 1999; Blanckenhorn et al. 2010), providing an upper limit for it because the magnitude of the temperature-size effect in the field is expected to be diluted to only a fraction thereof. I therefore suggest that temperature contributed to phenotypic body size variation in the field, in addition to other unspecified or unknown environmental factors, and conclude that the documented body size decrease in yellow dung flies over the past 15 years in Switzerland could relate to local (or even global) warming, though I definitely cannot prove it because the evidence presented here is correlational by nature. Dung limitation, deterioration, drying and/or competition in the dung could be alternative causes of the observed phenotypic body size decline, although a systematic change in these factors over the years is highly improbable given generally unpredictable spatio-temporal variation in cow dung (= larval food and adult oviposition substrate) availability and dung quality (Blanckenhorn 2009). Unfortunately, no quantitative field data are available to address these potentially confounding variables concerning the dung habitat. As already pointed out in the Methods, a systematic bias in measuring body size (hind tibia length) is also highly unlikely but can admittedly also not be excluded with absolute certainty.

Body size declines putatively related to global warming have been reported for some vertebrates (e.g. Yom-Tov and Greffen 2011; Van Buskirk et al. 2011 for North American birds; Caruso et al. 2014 for salamanders; reviewed by Merilä and Hendry 2014; Gotanda et al. 2015). For invertebrates, studies of allochronic (i.e. temporal) trait changes presumably driven by climate change are extremely scarce and have rarely been demonstrated to be genetic (see Bradshaw and Holzapfel 2001 for a nice example regarding diapause in the pitcher plant mosquito *Wyeomyia smithii*), recently prompting an entire special issue of the journal *Evolutionary Applications* on the subject (Merilä and Hendry 2014; Schilthuizen and Kellermann 2014; Stoks et al. 2014). The common-garden approach taken here shows that evolutionary responses of invertebrates in the field can be monitored, and strongly suggests that the documented body size changes of yellow dung flies are indeed genetic.

If anything, body size of yellow dung flies was rather expected to increase over the years, based on repeated assessments of predominant and strong sexual selection favoring large male size and fecundity selection favoring large female size in this (Jann et al. 2000; Kraushaar and Blanckenhorn 2002; Blanckenhorn et al. 2003; summarized in Blanckenhorn 2007; see Introduction) and many other species (Kingsolver and Pfennig 2004; Gotanda et al. 2015; e.g. Postma et al. 2007). In yellow dung flies there is also opposing juvenile viability selection against long development time, and hence indirectly against large body size because both are positively genetically correlated (Blanckenhorn 1999, 2002), which may be sufficient to balance size increases due to sexual selection under some but certainly not all environmental conditions (summarized in Blanckenhorn 2007; Teuschl et al. 2007). Few other disadvantages of large body size have so far been uncovered in this species. Demonstrations of counter-selection against large body size and/or of net selective advantages of small body size are generally difficult to come by (Blanckenhorn 2000; but see DiBattista et al. 2007; Carlson et al. 2008; Dufresne et al. 2009 for some recent counterexamples), to the extent that some researchers question the value of

phenotypic selection studies (Gotanda et al. 2015). Based on the total available evidence, therefore, evolution towards larger body sizes was *a priori* more reasonable to expect in yellow dung flies. Finding the opposite here is puzzling and difficult to explain. We must be missing important sources of selection against large body size, despite systematic search over the years (Blanckenhorn 2007, 2009). It may well be that the so far mechanistically not well understood physiological effects of warmer temperatures producing smaller bodies in ectotherms (the temperature-size rule: Atkinson and Sibly 1997) are stronger than the effects of net positive selection, thus indeed possibly masking the expected evolutionary response towards larger body sizes (Merilä et al. 2001). Alternatively, the empirical results of this study lend credence to the repeated argument that estimates of phenotypic selection in the wild, as well as laboratory heritability estimates, tend to be inflated, biased and/or overestimated, for various possible reasons (Merilä et al. 2001; Kingsolver and Pfennig 2004; Blanckenhorn 2007; Merilä and Hendry 2014; Gotanda et al. 2015). More long-term field studies of vertebrates and other invertebrates following the approach taken here should help resolve this question.

ACKNOWLEDGEMENTS

I thank the Swiss National Fund, the Zoological Museum and the University of Zurich for continued financial support, and multiple colleagues who have contributed to this research over the years (see publications listed in supplementary Table S2). Thanks also to Josh van Buskirk and Erik Postma for discussion, thought and comments on the manuscript.

LITERATURE CITED

Amano, K. 1983. Studies on the intraspecific competition in dung breeding flies: I. effects of larval density on yellow dung fly *Scatophaga stercoraria* L. (Diptera: Scathophagidae). Jap. J. Sanit. Zool. 34:165-175.

Andersson. M. 1994. *Sexual Selection*. Princeton University Press, Princeton, NJ.

Atkinson, D., and R.M. Sibly. 1997. Why are organisms usually bigger in cooler environments? Making sense of a life history puzzle. Trends Ecol. Evol. 12:235-239.

Blanckenhorn, W.U. 1997. Altitudinal life history variation in the dung flies *Scathophaga stercoraria* and *Sepsis cynipsea*. Oecologia 109:342-352.

Blanckenhorn, W.U. 1999. Different growth responses to temperature and resource limitation in three fly species with similar life histories. Evolutionary Ecology 13:395-409.

Blanckenhorn, W.U. 2000. The evolution of body size: what keeps organisms small? Quart. Rev. Biol. 75:385-407.

Blanckenhorn, W.U. 2002. The consistency of quantitative genetic estimates in field and laboratory in the yellow dung fly. Genetica 114:171-182.

Blanckenhorn, W.U. 2007. Case studies of the differential equilibrium hypothesis of sexual size dimorphism in dung flies. Pages 106-114 in: D.J. Fairbairn, W.U. Blanckenhorn, T. Székely, eds. SEX, SIZE AND GENDER ROLES. Evolutionary Studies of Sexual Size Dimorphism. Oxford University Press, Oxford.

Blanckenhorn, W.U. 2009. Causes and consequences of phenotypic plasticity in body size: the case of the yellow dung fly *Scathophaga stercoraria* (Diptera: Scathophagidae). Pages 369-422 in: D.W. Whitman, T.N. Ananthakrishnan, eds. Phenotypic plasticity of insects: Mechanism and consequences. Science Publishers, Enfield NH, USA.

Blanckenhorn, W.U., and M. Demont. 2004. Bergmann and converse Bergmann latitudinal clines in arthropods: two ends of a continuum? Int. Comp. Biol. 44:413-424.

Blanckenhorn, W.U., U. Kraushaar, and C. Reim. 2003. Sexual selection on morphological and physiological traits and fluctuating asymmetry in the yellow dung fly. J. Evol. Biol. 16:903-913.

Blanckenhorn, W.U., A.J. Pemberton, L.F. Bussière, J. Roembke, and K.D. Floate. 2010. Natural history and laboratory culture of the yellow dung fly, *Scathophaga stercoraria* (L.; Diptera: Scathophagidae). J. Insect Sci. 10:1-17.

Bradshaw, W.E., and C.M. Holzapfel. 2001. Genetic shift in photoperiodic response correlated with global warming. Proc. Nat. Acad. Sci. USA 98: 14509-14511.

Carlson, S.M., Olsen, E.M., and L.A. Vøllestad. 2008. Seasonal mortality and the effect on body size: a review and an empirical test using individual data on brown trout. Funct. Ecol. 22:663-673.

Caruso, N.M., Sears, M.W., Adams, D.C., and K.R. Lips. 2014. Widespread rapid reductions in body size of adult salamanders in response to climate change. Global Change Biol. 20:1751-1759.

Charmantier, A., L.E.B. Kruuk, J. Blondel, and M.M. Lambrechts. 2004. Testing for microevolution in body size in three blue tit populations. J. Evol. Biol. 17:732-743.

Charmantier, A., C. Perrins, R.H. McCleery, and B.C. Sheldon. 2006. Evolutionary response to selection on clutch size in a long-term study of the mute swan. Am. Nat. 167:453-465.

Chown, S.L., and K.J. Gaston. 2010. Body size variation in insects: A macroecological perspective. Biol. Rev. 85:139-169.

Coltman, D.W., J. Pilkington, L.E.B. Kruuk, A.J. Wilson, and J.M. Pemberton. 2001. Positive genetic correlation between parasite resistance and body size in a free-living ungulate population. Evolution 55:2116–2125.

DiBattista, J.D., K.A. Feldheim, S.H. Gruber, and A.P. Hendry. 2007. When bigger is not better: selection against large size, high condition and fast growth in juvenile lemon sharks. J. Evol. Biol. 20:201-212.

Dufresne, M., Lengfellner, K., and U. Sommer. 2009. Global warming benefits the small in aquatic ecosystems. Proc. Nat. Acad. Sci. USA 106:12788-12793.

Gotanda, K., Correa, C., Turcotte, M.M., Rolshausen, G., and A.P. Hendry. 2015. Linking macro-trends and micro-rates: re-evaluating micro-evolutionary support for Cope's rule. Evolution, in press.

Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Beerli, P., and J.G. Kingsolver. 2001. Strength and tempo of directional selection in the wild. PNAS 98:9157-9160.

Husby, A., S.M. Hille, and M.E. Visser. 2011. Testing mechanisms of Bergmann's rule: Phenotypic decline but no genetic change in body size in three Passerine bird populations. Am. Nat. 178:202-213.

Jann, P., W.U. Blanckenhorn, and P.I. Ward. 2000. Temporal and microspatial variation in the intensities of natural and sexual selection in the yellow dung fly *Scathophaga stercoraria*. J. Evol. Biol. 13:927-938.

Jensen, H., B.E. Saether, T.H. Ringsby, J. Tufto, S.C. Griffith, and H. Ellegren. 2003. Sexual variation in heritability and genetic correlations of morphological traits in house sparrows (*Passer domesticus*). J. Evol. Biol. 16:1296-1307.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P., and P. Beerli. 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157:245-261.

Kingsolver, J.G., and D.W. Pfennig. 2004. Individual-level selection as a cause of Cope's rule of phyletic size increase. Evolution 58:1608-1612.

Kraushaar, U., and W.U. Blanckenhorn. 2002. Population variation in sexual selection and its effect on body size allometry in two species of flies with contrasting sexual size dimorphism. Evolution 56:307-321.

Kruuk, L.E.B. 2004. Estimating genetic parameters in natural populations using the 'animal model'. Phil. Trans. Roy. Soc. Lond. B 359:873–890.

Merilä, J., and A.P. Hendry. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. Evol. Appl. 7: 1-14.

Merilä, J., B.C. Sheldon, and L.E.B. Kruuk. 2001. Explaining stasis: microevolutionary studies in natural populations. Genetica 112-113:199-222.

Mousseau, T.A., and D.A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59:181-197.

Nussey, D.H., A.J. Wilson, A. Morris, J.M. Pemberton, T.H. Clutton-Brock, and L.E.B. Kruuk. 2008. Testing for genetic trade-offs between early and late-live reproduction in a wild deer populations. Proc. Roy. Soc. Lond. B 275:745-750.

Postma, E., J. Visser, and A. Van Noordwijk. 2007. Strong artificial selection in the wild results in predicted small evolutionary change. J. Evol. Biol. 20:1823-1832.

Réale, D., M. Festa-Bianchet, and J.T. Jorgenson. 1999. Heritability of body mass varies with age and season in wild bighorn sheep. Heredity 83:526–532.

Réale, D., A.G. McAdam, S. Boutin, and D. Berteaux. 2003. Genetic and plastic responses of a northern mammal to climate change. Proc. Roy. Soc. Lond. B 270:591–596.

Roff, D.A. 1992. *The Evolution of Life Histories. Theory and Analysis*. Chapman and Hall, London, UK.

Roff, D.A. 1997. *Evolutionary Quantitative Genetics.* Chapman and Hall, London, UK.

Schilthuizen, M., and V. Kellermann. 2014. Contemporary climate change and terrestrial invertebrates: evolutionary versus plastic changes. Evol. Appl. 7:56–67.

Schluter, D., Price, T.D., and L. Rowe. 1991. Conflicting selection pressures and life history trade-offs. Proc. Roy. Soc. Lond. B. 246:11-17.

Stoks, R., Geerts, A.N., and L. de Meester. 2014. Evolutionary and plastic responses of freshwater invertebrates to climate change: realized patterns and future potential. Evol. Appl. 7:42-55.

Teuschl, Y., C. Reim, and W.U. Blanckenhorn. 2007. Correlated responses to artificial body size selection in growth development phenotypic plasticity and juvenile viability in yellow dung flies. J. Evol. Biol. 20:87-103.

Van Buskirk, J., R.S. Mulvihill, and R.C. Leberman. 2011. Declining body sizes in North American birds associated with climate change. Oikos 119:1047-1055.

Weigensberg, I., and D. A. Roff. 1996. Natural heritabilities: can they be reliably estimated in the laboratory? Evolution 50:2149-2157.

Wilson. A.J., J.M. Pemberton, J.G. Pilkington, T.H. Clutton-Brock, D.W. Coltman, and L.E.B. Kruuk. 2007. Quantitative genetics of growth and cryptic evolution of body size in an island population. Evol. Ecol. 21: 337-356.

Yom-Tov, Y., and E. Greffen. 2011. Recent spatial and temporal changes in body size of terrestrial vertebrates: probable causes and pitfalls. Biol. Rev. 86:531-541.

Figure Legend

Figure 1. Change in the average hind tibia length of (a) field-caught (N = 15 to 1101 per sample) and (b) laboratory-reared (N = 22 to 736 per sample) yellow dung fly males (squares and hatched line) and females (circles and solid line) over 15 years at Fehraltorf near Zurich. Note that lab-reared flies are considerably larger on average.

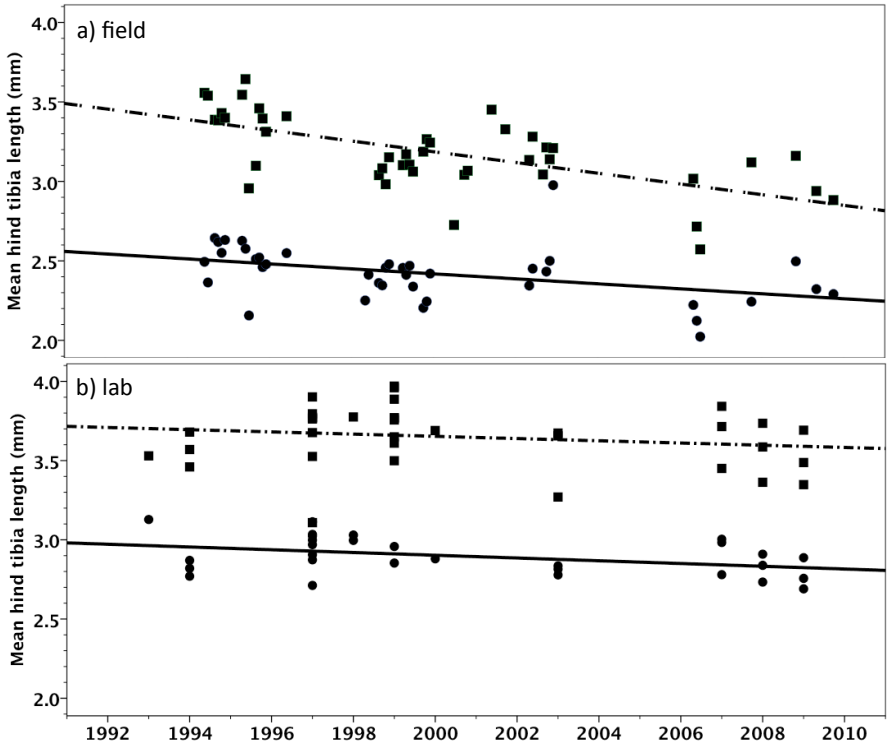

Table S1: Mean +/- SD (N) field sizes of female and male yellow dung flies in various years and months.

	female			male				Annual
Year.month	HTL (mm)	SD(m)	N(m)	HTL (mm)	SD(f)	N(f)	MeanTemp	Mean Temp
1994.05	2.49	0.34	125	3.56	0.33	525	14.25	10.6
1994.06	2.36	0.38	91	3.54	0.39	324	17.4	
1994.08	2.64	0.27	103	3.39	0.24	304	21.7	
1994.09	2.62	0.26	351	3.38	0.27	1101	13.9	
1994.10	2.55	0.26	304	3.43	0.26	911	10.8	
1994.11	2.63	0.22	212	3.40	0.27	674	8.9	
1995.04	2.63	0.25	129	3.55	0.29	422	10.75	9.3
1995.05	2.58	0.29	193	3.64	0.26	617	13.05	
1995.06	2.16	0.28	22	2.96	0.42	90	17.85	
1995.08	2.51	0.31	71	3.10	0.31	197	18.5	
1995.09	2.52	0.30	265	3.46	0.26	799	12	
1995.10	2.46	0.27	281	3.40	0.23	871	13.75	
1995.11	2.48	0.25	187	3.31	0.24	556	4.1	
1996.05	2.55	0.29	32	3.41	0.31	62	16.2	8.2
1997								9.4
1998.04	2.25	0.27	27				10.9	9.4
1998.05	2.41	0.26	116				15.6	
1998.08	2.36	0.31	18	3.04	0.31	97	19.85	
1998.09	2.35	0.24	69	3.08	0.25	383	15.45	
1998.10	2.46	0.28	61	2.98	0.28	403	11.5	
1998.11	2.48	0.23	46	3.15	0.25	171	0.9	
1999.03	2.46	0.31	15	3.10	0.30	66	8.2	9.4
1999.04	2.41	0.24	166	3.17	0.29	497	10.7	
1999.05	2.47	0.25	116	3.11	0.38	624	17.7	
1999.06	2.34	0.33	49	3.06	0.31	121	16.95	
1999.09	2.20	0.29	26	3.19	0.38	86	19.2	
1999.10	2.24	0.33	37	3.27	0.33	85	10.15	
1999.11	2.42	0.27	40	3.24	0.27	54	4.95	
2000.06				2.73	0.52	51	19.55	10.2
2000.09				3.04	0.28	126	16.65	
2000.10				3.07	0.24	131	10.4	
2001.05				3.45	0.24	81	17.15	9.5
2001.09				3.33	0.28	128	13.95	
2002.04	2.35	0.31	112	3.13	0.28	234	9.65	10.2
2002.05	2.45	0.28	96	3.28	0.35	242	15.85	
2002.08				3.04	0.31	28	20	
2002.09	2.43	0.28	38	3.21	0.27	88	13	
2002.10	2.50	0.20	40	3.14	0.23	86	11.6	
2002.11	2.78	0.27	54	3.21	0.26	60	9.25	
2003								10.0
2004								9.5
2005								9.3
2006.04	2.22	0.31	27	3.02	0.30	28	10.15	9.9
2006.05	2.12	0.24	29	2.72	0.24	29	13.85	
2006.06	2.02	0.23	32	2.57	0.31	31	16.9	
2007.09	2.24	0.25	141	3.12	0.28	125	14.75	10.3
2008.10	2.50	0.25	82	3.16	0.28	135	10.45	9.8
2009.04	2.32	0.22	61	2.94	0.36	39	12.8	9.8
2009.09	2.29	0.18	114	2.88	0.24	81	19.05	

Table S2: Mean +/- SD (N) laboratory sizes of female and male yellow dung flies at various temperatures over the years.

	female			male					
Year	HTL (mm)	SD(f)	N(f)	HTL (mm)	SD(m)	N(m)	Temp	Generation	Type
1993	3.13	0.10	117	3.53	0.09	328	20	F2	full-si
1994	2.87	0.09	80	3.68	0.09	79	15	F2	phen
1994	2.82	0.10	83	3.57	0.10	87	20	F2	phen
1994	2.77	0.09	82	3.46	0.10	88	25	F2	phen
1997	2.97	0.13	153	3.80	0.13	106	18	F1	phen
1997	3.11	0.08	35	3.90	0.12	22	18	F1	phen
1997	3.00	0.11	380	3.77	0.15	424	19	F2/F3	half-s
1997	3.03	0.09	254	3.76	0.10	351	21	F2/F3	half-s
1997	2.91	0.12	61	3.68	0.18	54	22	F1	phen
1997	3.02	0.11	22	3.77	0.06	28	22	F1	phen
1997	2.71	0.13	69	3.11	0.25	45	26	F1	phen
1997	2.87	0.13	45	3.53	0.15	22	26	F1	phen
1998	3.00	0.13	502	3.78	0.15	736	19	F2/F3	half-s
1998	3.03	0.12	220					F2/F3	half-s
1999				3.65	0.17	21	12	F2	phen
1999				3.96	0.11	25	15	F2	phen
1999				3.97	0.09	26	15	F2	phen
1999	2.96	0.12	189	3.76	0.16	170	15	F2	phen
1999				3.61	0.15	24	18	F2	phen
1999				3.89	0.17	60	19	F2	phen
1999				3.77	0.12	30	23	F2	phen
1999				3.77	0.13	31	23	F2	phen
1999	2.85	0.14	189	3.50	0.18	165	23	F2	phen
2000	2.88	0.11	152	3.69	0.11	258	19	F2/F3	full-si
2003	2.81	0.11	44	3.67	0.08	41	12	F2	full-si
2003	2.84	0.09	49	3.66	0.08	44	18	F2	full-si
2003	2.78	0.12	40	3.27	0.37	43	24	F2	full-si
2007	3.00	0.15	66	3.84	0.19	67	12	F2	full-si
2007	2.98	0.13	62	3.71	0.18	66	18	F2	full-si
2007	2.78	0.13	69	3.45	0.18	66	24	F2	full-si
2008	2.91	0.07	73	3.74	0.15	74	12	F2	full-si
2008	2.84	0.09	70	3.59	0.16	73	18	F2	full-si
2008	2.73	0.09	73	3.36	0.21	68	24	F2	full-si
2009	2.89	0.07	111	3.69	0.10	114	12	F2	full-si
2009	2.76	0.09	115	3.49	0.09	112	18	F2	full-si
2009	2.69	0.09	114	3.35	0.14	119	24	F2	full-si

REFERENCES:

Blanckenhorn WU (1997) Altitudinal life history variation in the dung flies *Scathophaga stercoraria* and *Sepsis cynipsea*. Oecologi

Blanckenhorn WU (2002) The consistency of heritability estimates in field and laboratory in the yellow dung fly. Genetica 114:171-

Blanckenhorn WU, Henseler C (2005) Temperature-dependent ovariole and testis maturation in the yellow dung fly. Entomologia

Blanckenhorn WU, Heyland A (2004) The quantitative genetics of two life history trade-offs in the yellow dung fly in abundant and

Blanckenhorn WU, Hosken DJ (2003) The heritability of three condition measures in the yellow dung fly. Behavioral Ecology 14: 6

Demont M, Blanckenhorn WU, Hosken DJ, Garner TWJ (2008) Population and quantitative genetic differentiation across Europe i

Hellriegel B, Blanckenhorn WU (2002) Environmental influences on the gametic investment of yellow dung fly males. Evolutionary

Hosken DJ, Blanckenhorn WU, Ward PI (2000) Developmental stability in yellow dung flies (*Scathophaga stercoraria*): fluctuating

926.

Jann P, Ward PI (1999) Maternal effects and their consequences for offspring fitness in the yellow dung fly. Functional Ecology 13

Mühlhäuser C, Blanckenhorn WU, Ward PI (1996) The genetic component of copula duration in the yellow dung fly, *Scathophaga*